
Developing a Fault Prediction Model for Wired Copper Networks under Precipitation

Abstract— Telecommunication companies who face challenges of aging infrastructure need to balance the cost of maintenance with that of providing their services within a service level guarantee. For Telstra, the largest telecommunication company in Australia, this balance is achieved by adopting a passive approach to handle the faults that occur in the network. Rather than actively preventing faults, technicians are assigned to fix faults in a timely manner. However, to achieve an efficient and timely technician assignment, a prediction model is needed to advise planners of the potential number of faults in the network. From statistical analysis, we have developed a fault prediction model by investigating 29 months of data of faults. Our prediction model shows that rain has a significant impact on the number of faults in many areas across Australia, which can be the result of the aging infrastructure.

Index Terms— Aging Infrastructure, Network Management, Fault Prediction Model, Statistical Analysis

1 INTRODUCTION

Wired networks are the backbone of communication networks all around the world and their efficient maintenance facilitates a sustainable growth of all dependent services. A fault prediction model enables the network operator to have early expectations of the location and number of faults and to prepare for treatment options. The preparations involve mobilizing equipment and technical staff to the locations where more faults are expected. A significant cost saving and improved maintenance planning are tangible outcomes of a fault prediction model. Weather as a primary source of faults in the wired networks is the focus of this study.

Wired networks can be mainly metal (copper) or optical fiber (glass). While fiber has a much higher capacity to transfer data, can transfer data in longer distances without repeaters, has a reasonable installation cost for new townships, and has a comparable wiring cost inside a building (Effenberger et al. 2007; Green 2004), copper still dominates the consumer market. In the OECD (Organization for Economic Cooperation and Development) area, only 22.1% of broadband subscriptions (average across member countries) were by fiber connection as of December 2016, although the market share is up to 70% in the leading countries of Japan and South Korea (OECD Broadband Portal 2016).

According to the Australian Bureau of Statistics report in December 2016, while the total number of internet subscribers (for ISPs with more than 1,000 subscribers) is 13.5 million, the share of fiber users is 10% and the share of DSL users (copper wire) is 35% (Australian Bureau of Statistics 2017). Mobile wireless has the greatest share with 45% of users.

According to another report from the Australian NBN Co, the number of fiber network users (both fiber to the premises, FTTP, and fiber to the node or neighborhood, FTTN) is reported to be 2 million as of June 2017 (NBN 2017). Out of the 2 million fiber users, 39% (957,919) are active end users on FTTN, who are still relying on the copper network. In such a combined infrastructure, the reliability of the fiber system is also dependent on the reliability of copper wires.

For years, it has been known among practitioners that weather conditions can impact copper wires more significantly than fiber cables. In addition, fiber networks have a smaller market share and are new which means they have less faults. This study, therefore, focuses on the copper wire network faults and how they correlate with weather conditions.

To perform our study, we have investigated a unique data set that includes 29 months data of the faults happening in the copper wired network of the largest telecommunication company in Australia, Telstra.

Telstra develops and maintains telecommunication networks of different types in the country. On a

daily basis, it receives tens of thousands of requests for its products and services, including installation of new connections, upgrade of systems, and regular or ad-hoc maintenance of network elements.

Many of the current services of Telstra are delivered over the Public Switched Telephone Network (PSTN) which is the traditional copper network. This network is being gradually phased out with the introduction of National Broadband Network (NBN) in Australia that delivers the required services mainly over optical fiber cables. However, due to the low penetration rate of the NBN, the majority of Telstra household and business customers still rely on the PSTN for receiving fixed telephone line, EFPTOS, Fax, and most importantly, ADSL services.

With aging of this infrastructure, the tradeoff between keeping the service level guarantee and the cost of the maintenance has become an issue for Telstra. To keep the balance, they follow an efficient resource allocation policy with the objective of assigning technicians to solve the faults reported by customers within a Customer Service Guarantee (CSG is a term used by the Australian Government to refer to a set of performance metrics that covers the supply of standard landline telephone services). In other words, in lieu of preventing faults from happening which requires investment in monitoring and maintaining the network and complex fault correlation and prediction analysis, they have a reactive strategy to fault handling.

Such a passive policy toward fault handling is based on the fact that the PSTNs are known to be very reliable despite their enormous size and complexity (Enriquez et al. 2002). One important factor in making these networks among the most reliable ones is identified as the possibility of quick human intervention (less than one hour) in recovery from failure (Kuhn 1997) which has been made possible due to a successful design based on loose coupling of system components. Such a design permits quick human intervention to recover from failure in most cases. This means that if the technicians with required skills are assigned to fix the faults in a timely manner, the fault can mostly be resolved within the CSG.

However, if the number of service failures reported by customers is considerably higher than the 'expected' volume, Telstra needs to move technicians from one state to another to meet the CSG requirements. The relocation of technicians is relatively costly, especially when it has not been provisioned in advance.

Therefore, Telstra would rather be proactive in this regard and foresee the significant changes in the pattern of faults, even on a short-term basis: if they can predict the need for extra resources in a timely manner, they can supply the required resources in other ways, for example by contracting them within the state, rather than moving resources from one state to another. Therefore, it is important for Telstra to know when there will be a significant change in the 'expected' trend of faults happening in the copper network.

Based on the insights from Telstra experts, rain is an important factor that can affect the expected pattern of faults. In fact, the planners who are responsible for managing the assignment of technicians to tasks perform their every day job while keeping an eye on the weather forecasts. However, this proposition has never been investigated, nor quantified.

This paper identifies the parameters affecting the demand for fixing services delivered over the Telstra PSTN; namely fixed telephone line and ADSL services. To the best of our knowledge, this study is the first to investigate the impact of rain on creating the faults in an aging PSTN. Furthermore, we have developed a prediction model based on the identified parameters.

To conduct this study, we were supplied with a unique data set that includes the records of fault reports and install jobs across Australia for 2 years and 5 months (29 months). This led us to adopt a statistical analysis approach as our research methodology. We developed the model by investigating various model structures. The final model is based on the analysis of covariance (ANCOVA) which combines analysis of variance (ANOVA) and regression models.

The remainder of this paper is organized as follows. Section 2 presents some background and related work, as well as highlights the contribution of our study. Section 3 elaborates on the statistical method of analysis in terms of the data, the method and the model. The process of data preparation and integration is explained in Section 4. The proposed prediction model is discussed in Section 5. Section 6 presents an extensive evaluation of the model and the discussion of the results. Future directions for the current study

are discussed in Section 7. Section 8 concludes the paper.

2 PRELIMINARIES

2.1 Short-term Planning

In this section, we further elaborate on the process of resource allocation in Telstra that allows a better understanding of the research question. This process is currently performed semi-automatically, in two steps: an automatic optimization process that is followed by a manual control process.

In the first step, a legacy system solves a complex optimization problem by considering the supply and demand: the available technicians and their skills on one hand, and the requests to fix a fault or install a new service, the required set of skills to respond to these requests, and the priority of the requests, on the other hand. The objective is to determine the optimal allocation of technicians to jobs for the next seven days across Australia. Due to the complexity of the optimization problem, it is solved at a specific geographic level, known as subdistrict level, once every 24 hours. As this optimization is performed at night, it is known as the *night optimizer*. A limited formulation of the optimization problem is solved for a subset of technicians in each subdistrict more frequently throughout the day. This is known as the *day optimizer*. Around 80% of the requests are handled successfully through this process.

In the second step, a manual monitoring and control process is performed which is known as the short-term planning with the objective of keeping a balance between supply and demand for the next three to four days. Planners are experienced people who have good knowledge of how the supply and demand may change in the subdistricts under their management, which subsequently can affect the optimality of the automated allocation. More specifically, they keep the balance by reviewing the assignments, handling the requests that have failed to get a technician during the optimization process (due to a set of predefined business constraints such as when there is a long distance between the technicians' base station and the fault's location), and finally close down the schedule of the next day for the subdistricts under their management.

Closing the assignments of a day prevents day and night optimizers from changing the assignment of that day or allocating more requests to it. The planners' decision to close down a day is based on the balance between available resources and scheduled tasks and their insight and experience of the subdistricts in terms of the actual performance of technicians, and the events that can impact the expected pattern of requests and/or the performance of the technicians, such as extreme weather conditions.

Moreover, they are responsible to initiate the process of asking for more technicians when the demand arises and when requests cannot be covered within a reasonable timeframe of the CSG.

The main challenge for planners is that the optimal assignment for the next seven days is based on the actual number of requests that have queued until the optimization time: it does not count for any potential variation in the service demand. This means that planners need to incorporate their personal insight and experience about the expected future demand and the events that can affect it to close the day assignments.

Particularly in case of the demand for fixing the faults, an unexpected change is not rare at all, mainly due to the aging infrastructure and its sensitivity to external events such as weather conditions. Therefore, if planners cannot foresee significant deviation of the service demand from its expected trend, they need to keep the balance between supply and demand by requesting technicians from other states. This is not desirable for the telecommunication company as relocating of technicians to areas outside their base can incur excessive costs.

2.2 Subdistrict: Request Management Level

To manage the copper network and services offered over this network, Telstra has divided Australia's PSTN to five regions: Northern, Central, Southern, SANT (South Australia and Northern Territory) and Western Australia. Each region is then divided to several districts, and each district is further divided to a number of subdistricts, with a total of 280 subdistricts across Australia. Short-term planning, as described above, is performed at a subdistrict level. Each planner is responsible for one or more subdistricts. This led us to the decision to perform the statistical analysis at the subdistrict level.

However, the network data (such as faults, installs, population of users) is collected and recorded at one

level below the subdistrict, known as the Exchange Service Area (ESA). The number of exchanges in a subdistrict varies and is a complex function of the parameters affecting the management of the network such as population, infrastructure, and supported services. There are more than 5000 exchanges in Australia. Therefore, as part of the data preparation process, we needed to define appropriate aggregation functions to aggregate data from an exchange level to a subdistrict level.

2.3 Related Work

In telecommunication systems, a variety of fault prediction models have been studied which can be divided into two broad categories based on the source of the problem: (1) faults caused by a problem in the software system; and, (2) faults occurring due to a failure of communication devices. Nevertheless, we should note that in complex scenarios, a fault might happen due to a mix of problems in software modules and network devices.

In the first category of prediction models, the objective is to predict which software modules are more likely to become faulty considering factors such as the size of the modules (Jung et al. 2004). A comprehensive review of the methods to develop fault prediction models in software systems is presented in (Arisholm et al. 2010). The second category of fault prediction models focuses on developing models based on alarms generated by network devices and collected by monitoring processes, such as the study performed in (Jaudet et al. 2005).

In the data set investigated in our study, when a customer reports a service disruption, a fault record is created and stored in the data set. At the time when the problem is reported, the source of the problem is generally unknown; it can be in the software system or network devices, or a mix of both. Triage is performed by technicians in the field. For us, the main research question was to investigate the impact of rain on such faults.

Based on the insight from Telstra planners and managers, rainfall is an important factor affecting the expected trend of the faults happening in the copper network. However, there is very little evidence in the literature supporting the relationship between rain and the number of faults in these networks.

Other than rain, some researchers have focused on extreme weather events and the potential impact these have on telecom networks. Weather-based disruptions (such as hurricanes and floods) in wired networks are studied in (Tornatore et al. 2016). They have suggested two high-level strategies for disaster-resiliency against these events, namely alert-based reconfiguration of the network connectivity and alert-based reconfiguration in cloud networks. To investigate the seasonality and periodicity of telecommunication network outages due to power loss, (Uchida 2014) has investigated the impact of extreme weather events on outages in the US. Their findings show that there are no seasonality effects in the outages, most likely due to the existence of backup power sources.

In electric power networks, studies in the US and Europe have shown that extreme weather events are the main cause of power outages (Chen et al. 2017). In these networks, precipitation-relevant parameters such as standardized precipitation index (SPI), and annual and monthly precipitation, have been studied in statistical models as some environmental factors contributing to the damage to these power networks (Wang et al. 2016).

In communication networks, the impact of precipitation on the performance of networks is mainly studied in the context of wireless communications. These studies include, for example, the impact of rain on wireless sensor networks (Ganesan et al. 2001), broadband wireless (Paraboni et al. 2002; Fong et al. 2004; Panagopoulos et al. 2007; Jabbar et al. 2008) and satellite communications (Harb et al. 2007; Luini and Capsoni 2013). In these networks, rain attenuation impacts the performance of the network where communication happens at carrier frequencies higher than 10GHz (Fong et al. 2004).

On the contrary to wireless communications and electric power networks, the studies to investigate the impact of rain on the wired networks in general, and Public Switched Telephone Networks (PSTN) in particular, are very limited.

The impact of extreme nature events on the US PSTN has been discussed in a study on the outages affecting more than 30,000 customers from April 1992 to March 1994 (Kuhn 1997). In this research, major sources of failure are identified as human error (from network company personnel and others), acts

of nature, and overloads with human errors contributing to 49% of the outages. An act of nature is defined as major and minor natural events that caused damage to the cable, power supply or facility, or natural disaster such as floods or hurricanes. A similar study was performed in 2000 on the US PSTN and they concluded that the share of human errors in causing failure has even increased more: from 49% in the first study to 54% in the second study. Although precipitation has not been mentioned explicitly in either of these studies, it can be categorized under major or minor natural events, especially when the infrastructure suffers from aging and is covered by a reactive (rather than a proactive) maintenance policy.

Aging infrastructure brings new challenges to these networks that were not considered as an issue before. The infrastructure is aging if new components are not added or the addition of new components is slower than the aging of existing components (Zuyi Li and Jiachun Guo 2006). Aging infrastructure faces challenges such as a higher cost to operate and maintain, and a lower reliability (Brown and Willis 2006). In many cases, the high cost of maintenance leads to poor maintenance decisions, such as relying on contractors who do not perform the jobs with acceptable quality. In case of the Telstra PSTN, aging has left the infrastructure vulnerable to rain (Ross 2012; Ross 2013).

To the best of our knowledge, this is the first study to investigate the impact of precipitation on the Public Switch Telephone Network in terms of the number of incurred faults. Moreover, as we had access to a unique data set, including around 2.5 years records of requests to fix faults and install new services, we were able to quantify this relationship with high accuracy across many areas in Australia, following statistical analysis techniques.

2.4 Contributions

This study led to a number of substantial outcomes. Firstly, we identify and elaborate on a set of parameters that can significantly impact the number of faults. More specifically, rainfall, number of installs, population of users, and the month of the year have significant relationships with the number of faults for many areas across Australia.

Secondly, we propose a statistical model to predict the number of faults based on the set of contributing parameters. The prediction power of the model differs from one area to another; however, in urban areas with sufficient rainy days, the accuracy of the prediction is high. Finally, we have evaluated the model in terms of its prediction power. We have presented the results Australia-wide, wherever possible. We have also drilled down the results for one specific area to give the reader a better view of the significance of the model and its prediction power.

3 A STATISTICAL ANALYSIS APPROACH

The main objective of this research is to investigate the relationship between the number of faults happening in the PSTN and the precipitation. Please note that, in general, precipitation refers to any form of water that falls to the earth's surface. However, in Australia, precipitation is mainly in form of rainfall. Therefore, in this paper, we have used rainfall and precipitation interchangeably.

If such a relationship exists between rainfall and number of faults, we aim to identify other potential variables contributing to the number of faults, relying on provided historical data set. The ultimate goal is to develop a model that can predict the number of faults based on the contributing variables.

3.1 Data

We were provided with a historical data set that includes (1) records of the faults reported by customers who experience a problem in their fixed telephone line or ADSL services, and (2) records of completed installation of fixed telephone line or ADSL services. The data was collected from across Australia for 29 months, from January 2013 to May 2015.

To investigate the main proposition, we also received historical precipitation data from the Australian Bureau of Meteorology (BOM 2017c) for 7,140 weather stations across Australia over the same time period. However, the examination of this data showed that the number of rainy days in 29 months of data can pose serious limitations on our analysis, due to not providing a statistically sufficient number of observations. To address this problem, we decided to consider the impact of rain in terms of rain classes, rather than the specific amount of rain. In this way, we can aggregate rainy days into meaningful classes,

with enough observations in each class, and then perform meaningful statistical inferences.

3.2 Method: Model Selection based on Historical Data

To perform the statistical analysis, we combined the two data sets (fault and install requests and precipitation data). The analysis has been undertaken through an iterative model selection process. We started with an exploratory data analysis, developed an initial model, and evaluated the model in terms of the statistical significance of the correlation between the dependent variable (number of faults) and the independent variables (such as precipitation), and also based on the prediction power of the model (the ratio of actual number of faults to the predicted number of faults). Each evaluation round led to improvements in the model by either applying a new statistical test, adding a new variable, or changing the type of a variable (for example from continuous to discrete).

The iterative exploratory statistical analysis led to identifying a set of contributing factors (independent variables) that can have a significant impact on the number of faults (the dependent variable). These factors are rain, number of installs, population of service users (PIO), and month of the year. PIO stands for “Pairs in Operation” which is an estimate of the population of users of Telstra’s wireline services.

Ideally, other contributing factors such as the age of infrastructure should have been included, but this data at the national level was not available. We investigated the impact of each variable on the number of faults, and also the potential interaction between some of the factors, as depicted in Table 1.

TABLE 1. POSSIBLE RELATIONSHIPS BETWEEN THE VARIABLES OF THE PREDICTION MODEL					
	Fault	Rain	Install	PIO	Month
Fault	-	In the model	In the model	In the model	In the model
Rain	-	-	Maybe, but not considered in the model*	-	Interaction in the model
Install	-	-	-	Maybe, but not considered in the model*	Maybe, but not considered in the model*
PIO	-	-	-	-	-
Month	-	-	-	-	-

*This interaction was not considered in the model as neither had we found any literature or theory supporting it, nor our initial data analysis provided evidence for it.

3.3 Model: Analysis of Covariance (ANCOVA)

As a general linear model, ANCOVA combines ANOVA and regression models. ANCOVA provides a way to control the (linear) effects of covariates on the dependent variable. Covariates are those continuous independent variables that are not the main aim of the research (e.g. number of installs and PIO in this study), however, they have a significant impact on the dependent variable. ANCOVA embeds the continuous independent variables in an ANOVA analysis which gives more statistical power to the analysis. ANCOVA considers the effects of covariates and then uses them in an ANOVA process. The effects of covariates are separated based on the levels of the factors as well as the combination of the factors and then their means are used for the prediction of the dependent variable.

Mathematically, ANCOVA decomposes the variance in the dependent variable (number of faults) into variance explained by the covariates (number of installs and PIO), variance explained by the categorical independent variable (classes of rain and month), and residual variance. ANCOVA uses the statistical technique (i.e. regression) to include the effects of the residual.

In this study, the dependent variable is the number of faults. This variable is supposed to be predicted by a constructed model based on the historical data. The main independent variable is the amount of rain. The occurrence of rain is low in Australia and the amount of rain significantly varies across different regions. Therefore, rainfall is classified to provide a basic level of required replicates (number of observations) for the analysis.

Our investigation shows that there is another factor that affects the average number of faults: seasons

or more specifically months. The impact of the month on the faults can be related to the rainfall or other parameters which cannot be inferred from the available dataset.

To study the relationship between rainfall and month, we investigated the interaction between these two factors. The results show that these two factors have some interactions but do not fully cover each other. In other words, we need to keep both variables in the model. This means that the impact of month on the number of faults is not solely related to the amount of rain in that month; rather, month *per se* has some impact on the number of faults. This made us change our analysis method from One-way ANOVA test to a factorial ANOVA.

The next step of the analysis shows that there are significant relationships between the number of faults and two other continuous variables: number of installs and PIO. Ignoring these two variables significantly decreases the accuracy of the prediction model. There are two options to take these variables into account in the constructed model. The first approach is to perform multiple regression analysis by taking the number of faults as the dependent variable, and the number of installations and PIO as independent variables. This approach requires the correlation between these two variables to be checked through some method such as stepwise regression. Then, the standard residual of faults should be used as the dependent variable in the factorial ANOVA. Standard residual of fault represents the linear effects of installs and PIO on the number of faults

The second approach is to use ANCOVA, where the number of install and PIO are added to the factorial model. Theoretically, the results of the two approaches are similar, so we decided to use ANCOVA. We also noted that the standard residual of the regression is perfectly normal, and it increases the accuracy of the ANOVA test. So, adding these two variables to the model effectively increases the accuracy of the prediction.

4 DATA PREPARATION AND INTEGRATION

In our study, data preparation for statistical analysis requires an estimate of the amount of rain received in each subdistrict, for each day of the analysis period. As the precipitation is measured at weather stations, a matching of weather stations to subdistricts has been done. However in doing so, we faced two problems: (1) not having a statistically sufficient number of rainy days during the 29 months of data, and (2) quality issues with the provided precipitation data. These two issues are addressed in the following sections.

4.1 Providing sufficient number of observations: local rain classification

To have statistically sufficient number of observations for the analysis, we decided to consider the impact of rain in terms of rain classes (categorical variable) rather than the exact amount of rain (continuous variable).

Initially, we designed a global rain classification for all the subdistricts; i.e. one classification for all. However, the results of the statistical analysis with global classification were not promising. This can be due to the fact that there are considerably large variations in the amount of rain and in the distribution of rainfall across regions in Australia. Therefore, we moved from the global one-size-fits-all rainfall classification to a more local rain classification for the subdistricts to enhance the power of the proposed prediction model.

To do so, subdistricts were grouped according to their climate, and a local rain classification was designed for each climate group. The grouping of subdistricts into climate classes is performed following a standard climate and rain categorization, suggested by the Australian Bureau of Meteorology, based on *seasonal rainfall*, illustrated in Fig. 1.

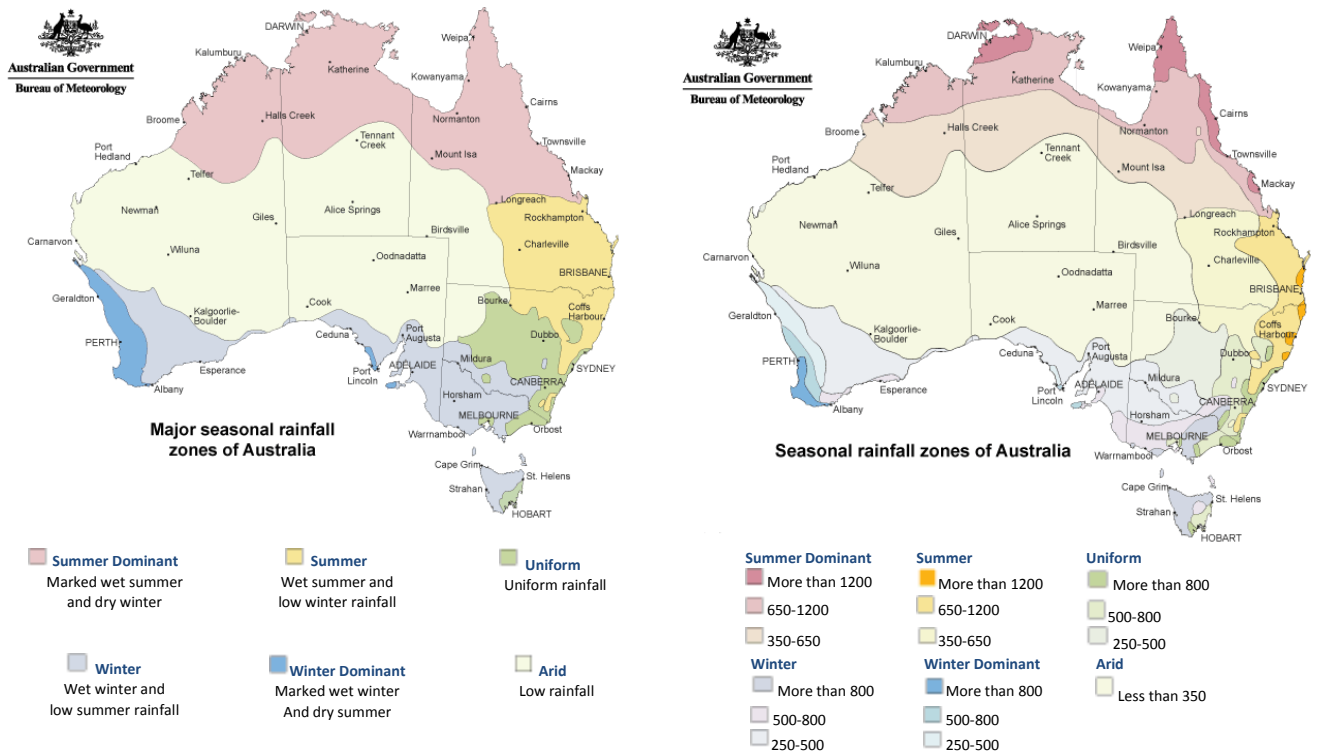


Fig. 1. Six major climate classes based on seasonal rainfall (BOM 2017a) Fig. 2. Australian climate classes based on seasonal rainfall (BOM 2017b)

In this categorization, seasonal rainfall is measured based on the median annual rainfall and seasonal incidence, where seasonal incidence is defined as the ratio of median rainfall received in November to April to the median of rainfall received in May to October.

The significance of this categorization is that it takes into account the median annual amount of rainfall over 100 years (1900-1999) rather than the average rainfall. Median rainfall offers a more representative metric for the distribution of rain throughout a year compared to the average. Moreover, the seasonal incidence improves the differentiation of the zones regarding the annual rain they receive.

As seen in Fig. 1, this categorization divides Australia to six major zones, namely summer dominant, summer, uniform, winter, winter dominant, and arid areas. The six categories indicate the season of highest rainfall in each zone.

A further classification of each major zone into three sub-zones is presented in Fig. 2, where 16 distinguished subzones are created based on numerical rainfall thresholds. We have used these 16 subzones to define the scope of the local rain classification. Hereafter, we refer to them as the *climate categories* in this paper.

We have developed a local rain classification for each climate category. The rain classifications may vary from one climate category to another, while all the subdistricts that fall into the same climate category have the same rain classification.

4.2 Quality issues of the precipitation data: dynamic assignment of weather stations to subdistricts

The precipitation data provided by the BOM has three important issues to consider. Firstly, according to BOM, "very few sites have a complete unbroken record of weather information. A site may have been closed, reopened, upgraded to a full weather site or downgraded to a rainfall only site during its existence causing breaks in the record for some or all elements." Secondly, the available precipitation data may not be acceptable after quality control; data may have been considered as wrong, suspected, or inconsistent. Finally, in some weather stations, the precipitation measurement is performed manually, which as a result, may not be measured on a daily basis. In these stations, the accumulated precipitation over the time between two measurements is reported.

Due to these issues, a fixed, static assignment of weather stations to subdistricts may lead to many days without valid rainfall information. To address this problem, we dynamically assigned weather stations to

subdistricts, on a daily basis.

The assignment of stations to subdistricts is performed in two steps. First, a list of candidate weather stations for each subdistrict is created. This is done through a GIS analysis process known as “Spatial Joint, Intersect”, using ArcGIS software. The output of this step is a prioritized list of stations for each subdistrict, where the stations within the subdistrict have the highest priority, and the others are ranked based on their distance from the border of the subdistrict. Then in the second step, on a daily basis, the candidate list is searched to find the weather stations which have quality-accepted data on a given day to create the dynamic list of assigned stations for each subdistrict.

5 THE PROPOSED PREDICTION MODEL

The contributing parameters to the number of faults in a subdistrict on a specific day are identified as:

- 1) The amount of rain (the classification of rain based on the climate category of the subdistrict)
- 2) The number of installs completed on that day
- 3) Month of the given date
- 4) PIO in that month

Accordingly, the proposed ANCOVA model to predict the number of faults based on the contributing variables, is formally presented in (1).

$$f_d^s = \beta_0^s + \beta_1^s \times t_d^s + \beta_2^s \times p_m^s + r_c^s + n_m^s + i_{c,m}^s \quad (1)$$

In this equation, f_d^s denotes the predicted number of faults on day d in subdistrict s , t_d^s denotes the number of completed installations on day d in subdistrict s , and p_m^s denotes the population of users (PIO) in month m in subdistrict s . A constant for each subdistrict is included, denoted as β_0^s . β_1^s and β_2^s are the coefficients for installs and PIO for each subdistrict, respectively. r_c^s denotes the adjusted value for rain classification c in subdistrict s . n_m^s denotes the adjusted value for month m in subdistrict s , and $i_{c,m}^s$ is the adjusted value for the interaction between rain category c and month m in subdistrict s .

The values of t_d^s and p_m^s can be obtained from the existing information systems in Telstra. β_0^s , β_1^s and β_2^s are calibrated for each subdistrict. The two contributing parameters, rain and month, are considered through their adjusted values for each subdistrict. Therefore, the values for these parameters vary from one subdistrict to another.

There is a possibility of correlation between the covariates as well as dependency of covariates to factors. For example, the demand for installs may vary across different months or may change with the changes in PIO. To evaluate the significance of these interactions, we evaluated the model for several subdistricts. The results showed that including these interactions did not improve the accuracy of the model significantly. Therefore, we considered the number of installs and PIO to be independent of rain.

6 RESULTS AND DISCUSSION

We have presented the results Australia-wide wherever possible. We have also drilled down the results for one district to give a better view of the significance of the model and its prediction power. The chosen subdistrict is known as “Northern Brisbane”, which is part of the Northern region in Australia.

Covering an area more than 51,000 square kilometres, this district includes Brisbane, the capital of Queensland, and major cities around it. Northern Brisbane includes 22 subdistricts that are different in terms of geographic size, population, being urban or remote, and climate category.

TABLE 2. CATEGORIZATION OF NORTHERN BRISBANE'S SUBDISTRICTS INTO LOW AND HIGH DENSITY (PIO/AREA)

Subdistrict Name	Climate Category*	Area (square kilometre)	PIO	Density	Category
CNB Ascot	[6]	85.37	68,820	806.09	HD
CNB Caboolture	[5]	5895.48	65,507	11.11	HD
CNB Caloundra	[5]	246.41	74,962	304.22	HD

CNB Chapel Hill	[6]	263.62	85,555	324.54	HD
CNB Cleveland	[6]	337.25	68,252	202.38	HD
CNB Coolumb Beach	[5]	453.55	33,547	73.97	HD
CNB Ipswich	[6]	789.57	52,002	65.86	HD
CNB Mitchelton	[6]	78.68	41,376	525.88	HD
CNB Mount Gravatt	[6]	153.05	75,162	491.10	HD
CNB Nambour	[5]	1881.39	29,118	15.48	HD
CNB Paddington	[6]	91.34	55,715	609.95	HD
CNB Petrie	[6]	480.76	44,326	92.20	HD
CNB Redcliffe	[5]	140.64	53,382	379.56	HD
CNB Woolloongabba	[6]	36.63	41,707	1138.50	HD
CNB Wynnum	[6]	120.34	58,762	488.31	HD
CNB Zillmere	[6]	82.16	47,275	575.39	HD
CNB Brisbane Islands	[5]	445.77	3,621	8.12	LD
CNB Dalby	[7]	11594.54	10,692	0.92	LD
CNB Gatton	[6]	3759.68	17,726	4.71	LD
CNB Gympie	[6]	4757.73	23,931	5.03	LD
CNB Kingaroy	[6]	11968.32	14,864	1.24	LD
CNB Toowoomba	[6]	7970.12	58,818	7.38	LD

*Climate categories 5,6,7 belong to the Summer class, where the amount of median annual rainfall for 5 is more than 1200mm, for 6 is in the range (650-1200)mm, and for 7 is in the range (350-650) mm.

To better demonstrate the results for this district, we have divided the map of this area into two parts based on the subdistricts' density; i.e. the ratio of the users' population (PIO) to area (square kilometre):

- High Density (HD) areas: density > 10,
- Low Density (LD) areas: density < 10 (
- Table 2).

6.1 Statistical Significance

In this section, we present the statistical significance of the ANCOVA model, which depicts the significance of the relationships between the dependent variable and independent variables. We have removed three remote islands from our statistical analysis, due to lack of enough historical rain data, namely Christmas Island, Coco Island and Thursday Island, making the number of subdistricts under study equal to 277.

The results of the ANCOVA analysis are presented in terms of the statistical significance of the relation between the main independent variable (rain category) and the dependent variable (number of faults) through the p-values as well as the R^2 of the model for each subdistrict.

The p-value of the rain category is presented in Fig. 3 as an intensity map. Studying this map leads to several interesting observations. Firstly, it shows that rain category and number of faults are significantly correlated for 92% of subdistricts (257 subdistricts). Secondly, the relationship is significant for all major cities and centres of population. However, drilling down the results shows that the CBD (central business district) of Brisbane, Melbourne, Sydney and Perth are exceptions to this observation: they are in populated urban areas with enough rain, but rain does not have significant relationship with the number of faults.

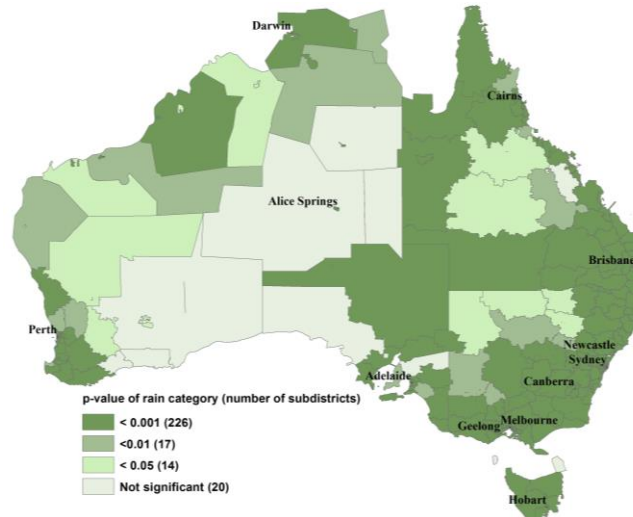


Fig. 3. The p-value of the relationship between rain category and number of faults

The reason is that in these subdistricts, the infrastructure is upgraded to NBN (new Australian broadband network) and therefore the service interruption is not sensitive to rain, unlike the copper network. There are other subdistricts that despite being in urban areas and having enough rain, the relationship is not significant, such as CNN Moranbah (near Brisbane) and CAS East Spencer (near Adelaide). Different reasons may have been contributing to such an impact, such as upgrading to the NBN, not having enough rain during the study time, or having many days with zero rain.

Thirdly, comparing this map with Fig. 1 shows that this relationship is not significant in arid areas in the centre and south of Australia with low rainfall. This can be due to several reasons, including the small size of the dataset for the rain classes, the skewness of the distribution of rainfall to zero amount of rain, and the lag in reporting the faults after rain.

We performed a similar study to investigate the statistical significance of the relationship between number of faults and number of installs, and also number of faults and PIO. The results are very similar to what is observed for the relationship of number of faults and rain: the relationship is statistically significant in urban areas, and not significant in remote, arid areas.

The p-values for all independent variables of subdistricts in Northern Brisbane are depicted in Table 3. As we can see, the relationship between rain category and number of faults, and the relationship between the number of installs and number of faults, are statistically significant for all subdistricts ($p < 0.001$). The p-value of the relationship between PIO and number of faults is also statistically significant for all subdistricts with $p < 0.001$, except for two subdistricts, Caloundra and Woolloongabba, where the relationship is not significant. Moreover, the relationship between month of the year and number of faults is significant with $p < 0.001$ for all subdistricts, except for one subdistrict (Ascot) where $p < 0.05$.

TABLE 3. THE P-VALUE OF THE RELATIONSHIP BETWEEN INDEPENDENT VARIABLES AND THE DEPENDENT VARIABLE (NUMBER OF FAULTS) FOR SUBDISTRICTS IN NORTHERN BRISBANE

Subdistrict Name	<i>p – value</i>			
	Rain Category	Number of Installs	PIO	Month
CNB Ascot	<0.001	<0.001	<0.001	0.011
CNB Caboolture	<0.001	<0.001	<0.001	<0.001
CNB Caloundra	<0.001	<0.001	0.76	<0.001
CNB Chapel Hill	<0.001	<0.001	<0.001	<0.001
CNB Cleveland	<0.001	<0.001	<0.001	<0.001
CNB Coolumb Beach	<0.001	<0.001	<0.001	<0.001
CNB Ipswich	<0.001	<0.001	<0.001	<0.001
CNB Mitchelton	<0.001	<0.001	<0.001	<0.001
CNB Mount Gravatt	<0.001	<0.001	<0.001	<0.001
CNB Nambour	<0.001	<0.001	<0.001	<0.001
CNB Paddington	<0.001	<0.001	<0.001	<0.001
CNB Petrie	<0.001	<0.001	<0.001	<0.001
CNB Redcliffe	<0.001	<0.001	<0.001	<0.001
CNB Woolloongabba	<0.001	<0.001	0.2	<0.001
CNB Wynnum	<0.001	<0.001	<0.001	<0.001

CNB Zillmere	<0.001	<0.001	<0.001	<0.001
CNB Brisbane Islands	<0.001	<0.001	<0.001	<0.001
CNB Dalby	<0.001	<0.001	<0.001	<0.001
CNB Gatton	<0.001	<0.001	<0.001	<0.001
CNB Gympie	<0.001	<0.001	<0.001	<0.001
CNB Kingaroy	<0.001	<0.001	<0.001	<0.001
CNB Toowoomba	<0.001	<0.001	<0.001	<0.001

6.2 R^2 : ANCOVA Analysis

The R^2 of the ANCOVA analysis, hereafter denoted as ANCOVA_ R^2 or R_a^2 , is presented in Fig. 4 as an intensity map. This map shows that the proposed model has a high prediction power ($R_a^2 > 0.45$) for all capital cities and centres of population. At the same time, when comparing this map with Fig. 2, we can see that these areas are grassland areas that enjoy median average rainfall above 500 mm.

The detailed information of the ANCOVA_ R^2 for Northern Brisbane is presented in Table 4. As we can see, the average R_a^2 across low density subdistricts is considerably below the overall average. While the average R_a^2 across high density subdistricts is above the overall average.

6.3 Actual Number of Faults versus the Predictions

In this section, we present the results of evaluating the prediction power of the model by comparing the actual number of faults (based on the available data set) against the predicted number of faults which are calculated following the below process.

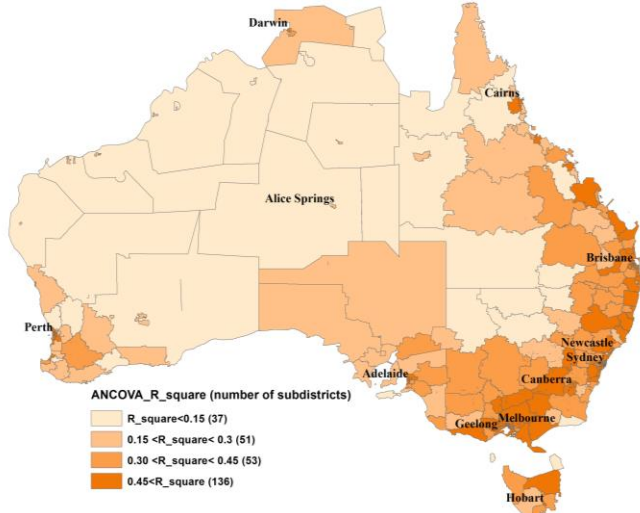


Fig. 4. Intensity map of the R^2 of the ANCOVA model for Australia

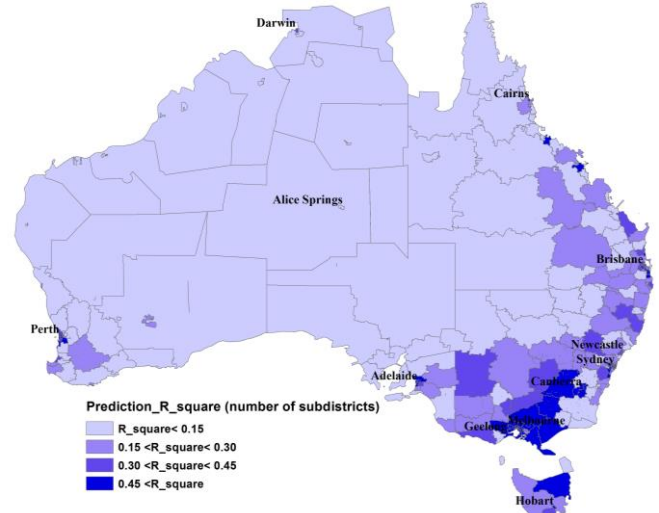


Fig. 5. The R^2 of the regressions analysis, presenting the prediction power of the model across Australia

For evaluating the model, we divided our data set into two parts: part one, using two years of data from 1/1/2013 to 31/12/2014, and part two, using five months of data from 1/1/2015 to 31/5/2015. We used part one as the training set, and calculated the coefficients of the model for each subdistrict based on this set. Part two was used as the test set: we predicted the number of faults based on the rain, number of installs, PIO and the coefficients' values derived from the training process. Then, we compared the predicted values and the actual values for these five months by performing a regression analysis. The R^2 of the regression, hereafter referred to as Prediction_ R^2 or R_p^2 , represents the prediction power of the model. A R_p^2 value of "1" shows that the predicted and actual faults are equal in all observations, while R_p^2 close to "0" shows that the model is not an adequate tool for prediction.

TABLE 4. THE ANCOVA_ R^2 FOR SUBDISTRICTS IN NORTHERN BRISBANE

	Subdistrict Name	ANCOVA_ R^2
High Density	CNB Ascot	0.69
	CNB Caboolture	0.63
	CNB Caloundra	0.66
	CNB Chapel Hill	0.67
	CNB Cleveland	0.54
	CNB Coolumb Beach	0.58
	CNB Ipswich	0.61

	CNB Mitchelton	0.64
	CNB Mount Gravatt	0.62
	CNB Nambour	0.49
	CNB Paddington	0.63
	CNB Petrie	0.58
	CNB Redcliffe	0.60
	CNB Woolloongabba	0.60
	CNB Wynnum	0.68
	CNB Zillmere	0.59
Low Density	Average	0.61
	CNB Brisbane Islands	0.19
	CNB Dalby	0.27
	CNB Gatton	0.44
	CNB Gympie	0.40
	CNB Kingaroy	0.34
	CNB Toowoomba	0.56
Overall Average		0.55

The results are presented as an intensity map in Fig. 5. The results show that for 72 subdistricts (about 30%), the Prediction_ R^2 is above 45%. Similar to what we observed about the ANCOVA_ R^2 , these subdistricts are urban areas, covering major cities which also have considerable annual rainfall. This means that the prediction power of the model is rather high in urban areas with enough service users and rainy days.

The detailed information of the Prediction_ R^2 for Northern Brisbane is presented in Table 5. We have calculated the average R_p^2 across high and low density areas. The high density areas enjoy more rain throughout the year, and they have higher user population which leads to a higher R_p^2 value on average across these subdistricts.

TABLE 5. THE PREDICTION_ R^2 FOR SUBDISTRICTS IN NORTHERN BRISBANE

	Subdistrict Name	Prediction_ R^2
High Density	CNB Ascot	0.61
	CNB Caboolture	0.30
	CNB Caloundra	0.56
	CNB Chapel Hill	0.53
	CNB Cleveland	0.28
	CNB Coolumb Beach	0.32
	CNB Ipswich	0.45
	CNB Mitchelton	0.35
	CNB Mount Gravatt	0.44
	CNB Nambour	0.33
	CNB Paddington	0.60
	CNB Petrie	0.33
	CNB Redcliffe	0.37
	CNB Woolloongabba	0.41
	CNB Wynnum	0.39
	CNB Zillmere	0.36
Low Density	Average	0.41
	CNB Brisbane Islands	0.05
	CNB Dalby	0.12
	CNB Gatton	0.22
	CNB Gympie	0.03
	CNB Kingaroy	0.12
	CNB Toowoomba	0.25
Overall Average		0.34

6.4 Applicability of the Model

Considering the ANCOVA_ R^2 (R_a^2) and Prediction_ R^2 (R_p^2), we are interested in the extent to which the model is useful. In other words, in how many subdistricts is the model calibrated and validated with high significance?

To answer this question, we defined the significance level of the model by simultaneous consideration of R_a^2 and R_p^2 which is illustrated in Table 6. In this table, the significance of the model is categorized into 4 levels, with the highest level of significance being 4, and the lowest level being 0. Level 4 is assigned to areas where both R_a^2 and R_p^2 are greater than 45%. Level 0 is assigned to areas where either of R_a^2 or

R_p^2 are below 15%.

TABLE 6. DEFINING THE SIGNIFICANCE LEVEL OF THE MODEL BY SIMULTANEOUS CONSIDERATION OF R_a^2 AND R_p^2

	$R_p^2 > 0.45$	$0.45 \geq R_p^2 > 0.3$	$0.3 \geq R_p^2 > 0.15$	$0.15 \geq R_p^2 \geq 0$
$R_a^2 > 0.45$	4	3	2	0
$0.45 \geq R_a^2 > 0.3$	3	2	1	0
$0.3 \geq R_a^2 > 0.15$	2	1	1	0
$0.15 \geq R_a^2 \geq 0$	0	0	0	0

R_a^2 : ANCOVA_ R^2 , R_p^2 : Prediction_ R^2

Following the definition and categorization in Table 6, the significance of the model is determined for all subdistricts across Australia and the result is plotted as an intensity map in Fig. 6.

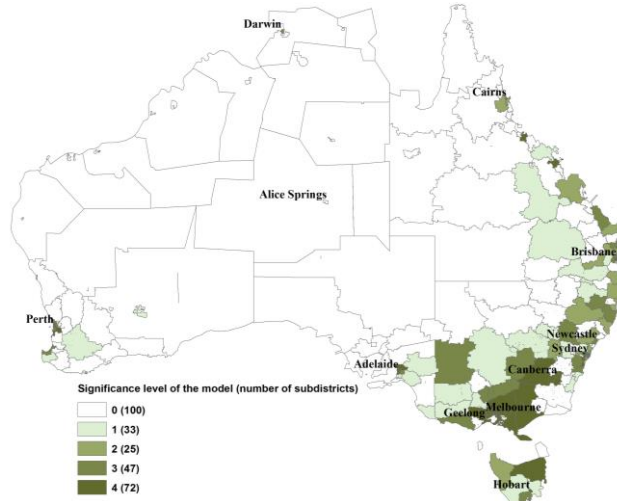


Fig. 6. Studying the applicability of the model through defining the significance level of the model

The results show that in urban areas with enough rain, the model can be very useful in predicting the number of faults. This includes 43% of subdistricts (47+72 out of 277) where both R_a^2 and R_p^2 are above 30%. However, in arid areas with neither much rain, nor considerable population, the model is not very useful. In these area, rain does not have a significant impact on the number of faults, and/or the prediction power of the model is relatively low. This include 36% of subdistricts, where any of R_a^2 or R_p^2 are below 15%.

6.5 More Users, Stronger Model

The statistical significance of the relationship between the independent variables and the dependent variable, and the prediction power of the model, vary across different subdistricts. However, considering Fig. 4, Fig. 5, and Fig. 6, it can be observed that the model has higher accuracy in urban areas where there are more telecommunication users. To formally test this observation, we performed a regression analysis to understand the relationships between ANCOVA_ R^2 and PIO, and that between Prediction_ R^2 and PIO.

The results of the regression of ANCOVA_ R^2 and PIO shows that for 277 observations, the R^2 is above 80%. Detail of the ANOVA test of this relationship is presented in Table 7. The R^2 of the regression of Prediction_ R^2 and PIO is over 78% for the same number of observations. Detail of the ANOVA test of this relationship is presented in Table 8.

TABLE 7. ANOVA RESULTS (ANCOVA_ R^2 AND PIO)

	<i>df</i>	<i>SS</i>	<i>Significance F</i>
Regression	1	9.269461	2.1E-99
Residual	275	2.255474	
Total	276	11.52494	

TABLE 8. ANOVA RESULTS (PREDICTION_ R^2 AND PIO)

	<i>df</i>	<i>SS</i>	<i>Significance F</i>
Regression	1	8.89416	3E-93
Residual	275	2.463838	
Total	276	11.358	

Plots of the line of best fit for the regression analysis in Fig. 7 also confirm our proposition that the

model accuracy is higher when the population of telecommunication users is higher. The increase in the accuracy of the model with increase in population can be due to having a higher number of service faults when there are more users. Having more valid data can lead to a better training and validation of the model. Another reason for such a relationship may be due to the fact that most of Australian cities and population are located in areas with higher annual rainfall. Similarly, having a larger sample of rainy days can lead to better training and validation of the model.

6.6 The Impact of Installs

Based on our study, the number of installs has a significant relationship with the number of faults. Accordingly, the number of installs is considered as one of the independent variables in the prediction model. Such a relationship can be justified with regard to the literature on the impact of human errors on the faults happening in PSTN networks.

In the study performed on the US PSTN (Kuhn 1997), errors on the part of telephone company personnel and acts of nature caused similar amounts of downtime (14 and 18 percent). The number of outages caused by human error from the company's personnel was 25%. A later study showed an increase in the role of the telephone company's personnel on the number of outages, pertaining to 38% of the outages in the USA (Enriquez et al. 2002).

We speculate that the relationship of the number of installs and the number of faults is due to a similar reason: it is likely that when personnel install a new service, they may damage the cable, power supply or other parts of the facility.

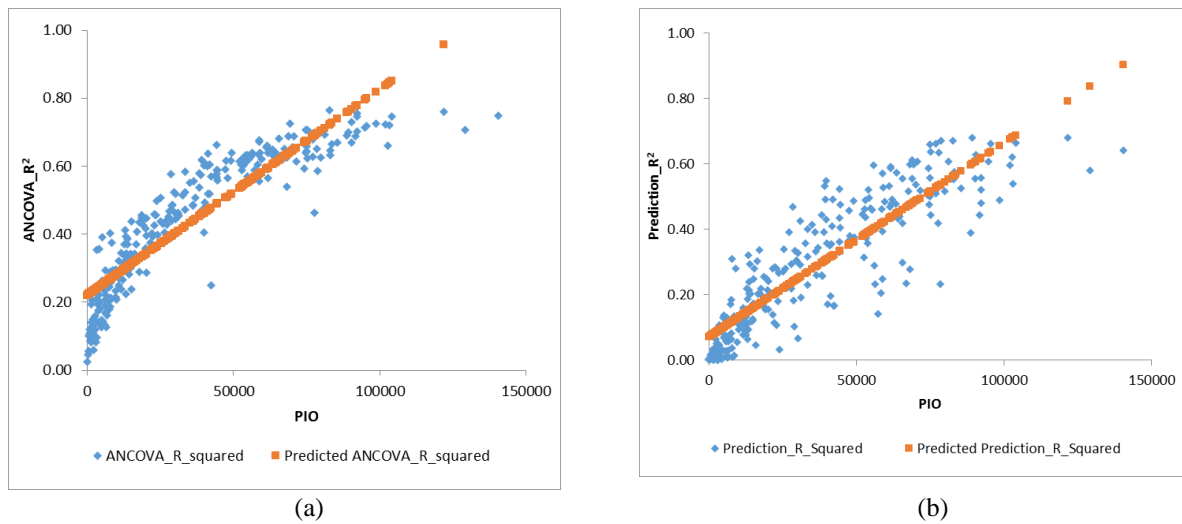


Fig. 7. Plotting the line of best fit for regression analysis of (a) ANCOVA_R² and (b) Prediction_R² against PIO

7 FUTURE DIRECTIONS

The results of this study should be considered within certain limits for practical applications. The following limitations should be considered for future model development.

7.1 Size of the Data

This study is performed based on historical fault data for 29 months. As there are not many rainy days throughout the year in many parts of Australia, there may not have been enough observations (faults and rain). This may create a bias toward the many days of not having any rain, thus possibly under-predicting the number of faults on rainy days. Performing the analysis for longer period of time has the potential to improve the prediction power of the model significantly.

7.2 Method of Data Collection

Not all desirable independent variables were available for this study. Therefore, we had to rely on existing, historical data that is not specifically designed for this study. In other words, we were not able to consider some of the potential independent variables that may affect the number of faults, such as the age of the infrastructure or the type of users (household/business).

7.3 Weather Conditions

In this study, we did not differentiate between different weather events, mainly due to lack of data. While some faults may be happening due to rainfall, other weather conditions such as storms, wind, extreme heat and lightning can be important causes of faults in a wireline network.

7.4 Rainfall Data

We have based our analysis on the maximum rainfall observed from three weather stations, preferably inside a subdistrict. However, for some of the subdistricts, for some of the days, we had to rely on weather information from outside the subdistrict or had data of one weather station only.

7.5 Analysis at the Subdistrict Level

The analysis is performed at the subdistrict level, while the fault data is available at the exchange level. Therefore, we required heuristics for aggregate factors, such as climate category, at the subdistrict level. However, for some factors such as Customer Service Guarantee (CSG), there may not be a meaningful aggregation function. This prevented us from including the CSG category as a potential contributing factor in our analysis.

Moreover, subdistricts significantly vary in size from one another. For the larger ones, it is particularly difficult to have only one meaningful value as a representative of the rainfall in that subdistrict.

The prediction model can be further improved if the analysis is performed at the exchange level, and then exchanges are grouped into meaningful sets with regards to their pattern of fault occurrence, considering factors such as CSG.

8 CONCLUSION

In this paper, we have studied an important problem in aging telecommunication networks: the impact of rain on the faults in the network. More specifically, we have studied the faults interrupting fixed telephone line and ADSL services in the Australia PSTN. To conduct our study, we initially identified the variables affecting the number of faults in this network, which are rainfall, installations and the population of the service users. We also proposed a model that can predict the number of faults based on these variables.

For our study, we had access to a unique data set of 29 months of fault and installation records related to fixed telephone and ADSL services across Australia. Accordingly, this study has followed statistical methods. We studied the statistical significance of the relationship between the identified parameters (rain, installs, and user's population) and the number of faults. Our studies show that the relationship is significant in all capital cities and centres of population.

We performed extensive evaluation on the proposed model. For evaluation, the data set was divided to (1) two years of data for training of the model, and (2) five months data for testing the model. We evaluated the model in terms of its prediction power and the applicability of the model. Based on the evaluation results, the prediction power of the model is higher in urban areas. A simultaneous analysis of (1) the statistical significance of the relationship of independent variables and the dependent variable, and (2) the prediction power of the model, shows that the model is most useful in urban areas which enjoy rainfall throughout the year.

To the best of our knowledge, this study is the first to consider the impact of rainfall on aging PSTNs. Moreover, it is the first study to develop a quantitative measure for the number of faults happening in such networks. The results of our study are of significant value for researchers and practitioners in telecommunication network and infrastructure domain for making decisions regarding the management and maintenance of these networks.

REFERENCES

- Arisholm, E., Briand, L.C. & Johannessen, E.B., 2010. A systematic and comprehensive investigation of methods to build and evaluate fault prediction models. *Journal of Systems and Software*, 83(1), pp.2–17.
- Australian Bureau of Statistics, 2017. *Internet Activity, Australia, December 2016, Main Features - Type of access*

- connection, Commonwealth of Australia; Australian Bureau of Statistics.
- BOM, 2017a. Australian Climate Averages - Climate classifications-Köppen - major classes. Available at: http://www.bom.gov.au/jsp/ncc/climate_averages/climate-classifications/index.jsp?maptype=kpng#maps [Accessed November 20, 2017].
- BOM, 2017b. Australian Climate Averages - Climate classifications - Köppen - all classes. Available at: http://www.bom.gov.au/jsp/ncc/climate_averages/climate-classifications/index.jsp?maptype=kpn#maps [Accessed November 20, 2017].
- BOM, 2017c. Australian Government, Bureau of Meteorology. Available at: <http://www.bom.gov.au/climate/data/> [Accessed November 20, 2017].
- Brown, R.E. & Willis, H.L., 2006. The economics of aging infrastructure. *IEEE Power and Energy Magazine*, 4(3), pp.36–43.
- Chen, C., Wang, J. & Ton, D., 2017. Modernizing Distribution System Restoration to Achieve Grid Resiliency Against Extreme Weather Events: An Integrated Solution. *Proceedings of the IEEE*, pp.1–22.
- Effenberger, F., Cleary, D., Haran, O., Kramer, G., Li, R., Oron, M. & Pfeiffer, T., 2007. An introduction to PON technologies [Topics in Optical Communications]. *IEEE Communications Magazine*, 45(3), pp.S17–S25.
- Enriquez, P., Brown, A. & Patterson, D.A., 2002. Lessons from the PSTN for dependable computing. In *Workshop on Self-Healing, Adaptive and Self-Managed Systems*.
- Fong, B., Ansari, N., Fong, A.C.M. & Hong, G.Y., 2004. On the scalability of fixed broadband wireless access network deployment. *IEEE Communications Magazine*, 42(9), pp.S12–S18.
- Ganesan, D., Govindan, R., Shenker, S. & Estrin, D., 2001. Highly-resilient, energy-efficient multipath routing in wireless sensor networks. *ACM SIGMOBILE Mobile Computing and Communications Review*, 5(4), p.11.
- Green, P.E., 2004. Fiber to the home: the next big broadband thing. *IEEE Communications Magazine*, 42(9), pp.100–106.
- Harb, K., Srinivasan, A., Huang, C. & Cheng, B., 2007. QoS In Weather Impacted Satellite Networks. In *IEEE Pacific Rim Conference on Communications, Computers and Signal Processing*. IEEE, pp. 178–181.
- Jabbar, A., Raman, B., Frost, V.S. & Sterbenz, J.P.G., 2008. Weather disruption-tolerant self-optimising millimeter mesh networks. In K. A. Hummel & J. P. G. Sterbenz, eds. *Self-Organizing Systems*. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, pp. 242–255.
- Jaudet, M., Lqbal, N., Hussain, A. & Sharif, K., 2005. Temporal classification for fault-prediction in a real-world telecommunications network. In *Proceedings of the IEEE Symposium on Emerging Technologies*. IEEE, pp. 209–214.
- Jung, H.-W., Lim, Y. & Chung, C.-S., 2004. Modeling change requests due to faults in a large-scale telecommunication system. *Journal of Systems and Software*, 72(2), pp.235–247.
- Kuhn, D.R., 1997. Sources of failure in the public switched telephone network. *Computer*, 30(4), pp.31–36.
- Luini, L. & Capsoni, C., 2013. On the Relationship Between the Spatial Correlation of Point Rain Rate and of Rain Attenuation on Earth-Space Radio Links. *IEEE Transactions on Antennas and Propagation*, 61(10), pp.5255–5263.
- NBN, 2017. *NBN annual results presentation 2016-2017*, Australia.
- OECD Broadband Portal, 2016. Percentage of fibre connections in total broadband (Dec. 2016). *The Organisation for Economic Co-operation and Development (OECD)*. Available at: <http://www.oecd.org/sti/broadband/oecdbroadbandportal.htm> [Accessed September 7, 2017].
- Panagopoulos, A.D., Arapoglou, P.-D.M., Kanellopoulos, J.D. & Cottis, P.G., 2007. Intercell Radio Interference Studies in Broadband Wireless Access Networks. *IEEE Transactions on Vehicular Technology*, 56(1), pp.3–12.
- Paraboni, A., Masini, G. & Elia, A., 2002. The effect of precipitation on microwave LMDS networks-performance analysis using a physical raincell model. *IEEE Journal on Selected Areas in Communications*, 20(3), pp.615–619.
- Ross, N., 2012. Australia's copper network is "Rooted." *ABC TECHNOLOGY AND GAMES*. Available at: <http://www.abc.net.au/technology/articles/2012/11/23/3639761.htm> [Accessed February 15, 2016].
- Ross, N., 2013. NBN alternative: Is Australia's copper network fit for purpose? *ABC TECHNOLOGY AND GAMES*. Available at: <http://www.abc.net.au/technology/articles/2013/09/19/3851924.htm> [Accessed February 15, 2016].
- Tornatore, M., Andre, J., Babarczy, P., Braun, T., Folstad, E., Heegaard, P., Hmaity, A., Furdek, M., Jorge, L., Kmiecik, W., Mas Machuca, C., Martins, L., Medeiros, C., Musumeci, F., Pasic, A., Rak, J., Simpson, S., ... Voyiatzis, A., 2016. A survey on network resiliency methodologies against weather-based disruptions. In *8th International Workshop on Resilient Networks Design and Modeling (RNDM)*. IEEE, pp. 23–34.
- Uchida, M., 2014. Statistical characteristics of serious network failures in Japan. *Reliability Engineering & System Safety*, 131, pp.126–134.
- Wang, Y., Chen, C., Wang, J. & Baldick, R., 2016. Research on Resilience of Power Systems Under Natural Disasters—A Review. *IEEE Transactions on Power Systems*, 31(2), pp.1604–1613.
- Zuyi Li & Jiachun Guo, 2006. Wisdom about age [aging electricity infrastructure. *IEEE Power and Energy Magazine*, 4(3), pp.44–51.